

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2000-322061
(P2000-322061A)

(43) 公開日 平成12年11月24日 (2000.11.24)

(51) Int.Cl.	識別記号	F I	テームト* (参考)
G 1 0 H 1/043		G 1 0 H 1/043	Z 5 D 0 4 5
G 1 0 L 21/04		G 1 0 L 3/02	C 5 D 3 7 8

審査請求 有 請求項の数 4 O L (全 8 頁)

(21) 出願番号 特願平11-126349
(22) 出願日 平成11年5月6日 (1999.5.6)

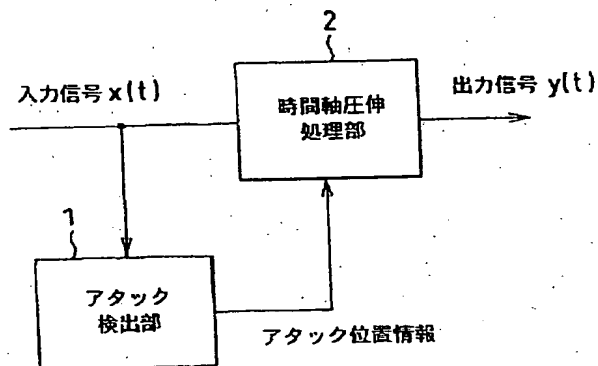
(71) 出願人 000004075
ヤマハ株式会社
静岡県浜松市中沢町10番1号
(72) 発明者 近藤 多伸
静岡県浜松市中沢町10番1号 ヤマハ株式
会社内
(74) 代理人 100092820
弁理士 伊丹 勝
Fターム (参考) 5D045 BB01
5D378 AD07 BB12 HB11 HB13 MM27

(54) 【発明の名称】 リズム音源信号の時間軸圧伸方法及び装置

(57) 【要約】

【課題】 リズム音源信号に対し時間軸圧伸処理を施す際に、時間軸圧伸後のアタック位置を補償して、リズムの狂いや二度打ちを発生させることを防止する。

【解決手段】 リズム音源の入力信号 $x(t)$ は、アタック検出部1に入力され、ここでアタック位置を検出される。入力信号 $x(t)$ は、時間軸圧伸処理部2にも供給されている。時間軸圧伸処理部2は、入力信号 $x(t)$ のうち、アタック検出部1で検出されたアタック位置間の信号について時間軸圧伸処理を施す。



【特許請求の範囲】

【請求項1】 時間軸圧伸処理すべきリズム音源信号からアタック位置を検出し、この検出されたアタック位置の間のリズム音源信号に対して時間軸圧伸処理を施すようにしたことを特徴とするリズム音源信号の時間軸圧伸方法。

【請求項2】 前記リズム音源信号のうち、前記検出されたアタック位置とその近傍とを除いた部分について時間軸圧伸処理を行い、この時間軸圧伸処理された信号の両端を時間軸圧伸処理されない信号と滑らかに結合するようにしたことを特徴とする請求項1記載のリズム音源信号の時間軸圧伸方法。

【請求項3】 時間軸圧伸処理すべきリズム音源信号からアタック位置を検出するアタック位置検出手段と、このアタック位置検出手段で検出されたアタック位置間のリズム音源信号をピッチを変えずに予め指定された圧伸率で時間軸圧伸処理する時間軸圧伸処理手段とを備えたことを特徴とするリズム音源信号の時間軸圧伸装置。

【請求項4】 前記時間軸圧伸処理手段は、前記リズム音源信号のうち、前記検出されたアタック位置とその近傍とを除いた部分について時間軸圧伸処理を施し、この時間軸圧伸処理された信号の両端を時間軸圧伸処理されない信号と滑らかに結合するものであることを特徴とする請求項3記載のリズム音源信号の時間軸圧伸装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、原デジタル信号のピッチを変えずに原デジタル信号を所望とする圧伸率で時間軸圧伸するデジタル信号の時間軸圧伸方法及び装置に関し、特にリズム音源信号に対する時間軸圧伸方法及び装置に関する。

【0002】

【従来の技術】デジタル・オーディオ信号のピッチを変えずにその時間軸を圧縮又は伸長する時間軸圧伸技術は、例えば、収録されたデジタルオーディオ信号の全体的な収録時間を所定の時間に合わせ込む、いわゆる「尺合わせ」や、カラオケ装置等のテンポ変換等に利用される。この種の時間軸圧伸技術としては、従来より、カット・アンド・スプライス法（例えば特開平10-282963号）やポインター移動量制御による重複加算法（“ポインター移動量制御による重複加算法を用いた音声の時間積での伸長圧縮とその評価”；森田、板倉、昭和61年10月；日本音響学会秋期大会講演論文集1-4-14, PP149）等が知られている。

【0003】一般的なカット・アンド・スプライス法による時間軸圧伸処理は、波形とは無相関に波形の切り出しを行い、切り出した波形をつなぎ合わせて指定された圧伸率での圧伸処理を行うもので、切り出し波形同士のつなぎの部分では、波形の不連続が生じるので、クロスフェード処理を行って、フレームのつなぎ部分を滑らか

にする。切り出し間隔は、人間の聴感上、エコー感や音のダブリ感が検知されにくい間隔、例えば60msec程度に設定される。特に、特開平10-282963号の方式では、音声タイミング情報に同期して切り出しの長さを決める。通常の方式に比べ、元波形のリズムと同じ周期でつなぎ目が現れ、つなぎ目の部分の音質変化が目立ちにくいという特徴がある。

【0004】一方、ポインター移動量制御による重複加算法では、原オーディオ信号において、波形相関の最も高い隣接した同じ長さの2つの区間を抽出し、これら区間の信号を重複加算し、この重複加算された信号を元の2つの区間と入れ換えたり、元の2つの区間の間に挿入することで、全体的な時間を短くしたり長くする。この方式は、カット・アンド・スプライス法よりもスムーズな波形接続が可能となる。特に音声信号や単音楽器のようなピッチ性の高い音源に対し、より品質の高い時間軸圧伸が可能となる。

【0005】

【発明が解決しようとする課題】従来の一般的なカット・アンド・スプライス法では、どのような信号を対象としても、それなりの音質が期待できるというメリットはあるものの、特にリズム音源を対象とした場合には、二度打ちやリズムの狂いといった非常に目立つ音質劣化を発生させやすい。特開平10-282963号の方式では、元波形のリズムに同期したカット・アンド・スプライスとなっているが、特に伸長の場合、波形を切り出す際に2つのアタックが切り出し波形に含まれることがあり、この場合二度打ちが発生する。更に、ポインタ移動量制御による重複加算法では、波形の時間相関を見ながら時間軸圧伸を行うため、二度打ち現象は原理的には起きないと考えられる。しかし、時間軸圧伸後のアタックの位置については全く補償されておらず、この結果、リズムのずれが発生し易い。

【0006】この発明は、このような問題点を鑑みなされたもので、リズム音源信号に対し時間軸圧伸処理を施す際に、時間軸圧伸後のアタック位置を補償して、リズムの狂いや二度打ちを発生させることがないリズム音源信号の時間軸圧伸方法及び装置を提供することを目的とする。

【0007】

【課題を解決するための手段】この発明に係るリズム音源信号の時間軸圧伸方法は、時間軸圧伸処理すべきリズム音源信号からアタック位置を検出し、この検出されたアタック位置の間のリズム音源信号に対して時間軸圧伸処理を施すようにしたことを特徴とする。

【0008】また、この発明に係るリズム音源信号の時間軸圧伸装置は、時間軸圧伸処理すべきリズム音源信号からアタック位置を検出するアタック位置検出手段と、このアタック位置検出手段で検出されたアタック位置間のリズム音源信号をピッチを変えずに予め指定された圧

伸率で時間軸圧伸処理する時間軸圧伸処理手段とを備えたことを特徴とする。

【0009】この発明によれば、リズム音源信号のアタック位置を検出し、検出されたアタック位置間で時間軸圧伸処理を施すようにしているので、二度打ちが発生することはなく、アタック位置の間隔も圧伸率に応じて圧縮又は伸長されることになり、これにより圧伸処理前後のアタック位置の相対関係は完全に維持されるので、リズムの狂いが発生するようなこともない。

【0010】リズム音源信号のうち、検出されたアタック位置とその近傍とを除いた部分について時間軸圧伸処理を行い、この時間軸圧伸処理された信号の両端を時間軸圧伸処理されない信号と滑らかに結合するようにする。滑らかに結合させるには、時間軸圧伸処理の際に、両端部での処理波形が元の信号波形とほぼ似通うようにしたり、或いはクロスフェード処理で結合させる。この場合、アタックの部分の波形はそのまま維持されるので、本来の音に近い音が得られる。

【0011】

【発明の実施の形態】以下、図面を参照して、この発明の好ましい実施の形態について説明する。図1は、この発明の一実施例に係るリズム音源信号の時間軸圧伸装置の構成を示すブロック図である。時間軸圧伸すべきリズム音源信号であるディジタル・オーディオ信号 $x(t)$ は、アタック検出部1に入力されている。即ち、リズム音源信号には、“アタック”が存在し、波形レベルでは信号電力の急激な集中と変化となっている。アタック検出部1では、あるしきい値によって単位時間当たりの信号電力の評価を行うと共に、この信号電力の時間微分によって、波形の急激な変化点を検出する。この2つを組み合わせてることにより、リズム音源内のほぼ全てのアタックの検出が可能である。そして、この検出結果は、アタック位置情報として出力される。

【0012】一方、入力オーディオ信号は、時間軸圧伸処理部2にも供給されている。時間軸圧伸処理部2は、入力されたオーディオ信号のうち、アタック検出部1で検出されたアタック位置間の信号について時間軸圧伸処理を施す。この時間軸圧伸処理部2における圧伸方式としては、カット・アンド・スプライス法、ポインタ移動量制御による重複加算法、リバブ、ディザ、ループの繰り返し等、種々の方法を適用することができるが、ここでは、一例としてポインタ移動利用制御による重複加算法による圧伸方式について説明する。

【0013】図2は、時間軸圧伸処理部2の構成を示すブロック図である。入力オーディオ信号は、遅延バッファ11に順次格納される。遅延バッファ11は、波形の時間軸伸長処理およびピッチ抽出処理等に必要なデータ量が格納されるリングバッファである。遅延バッファ11に格納されたオーディオ信号は、隣接波形読出位置制御部12の制御に基づき種々の区間長で切り出され、隣

接波形のデータとして順次読み出される。類似度計算部13は、隣接波形読出制御部12の制御のもとで読み出された隣接波形のデータの類似度を計算する。制御部14は、求められた類似度から隣接波形が最も類似する区間長を求め、これを基本周期（ピッチ） L_p として波形読出部15に与える。波形読出部15は、制御部14に与えられた前記アタック位置情報に基づき、アタック間の信号について、与えられた基本周期 L_p だけ離れた2つのデータを遅延バッファ11から読み出す。遅延バッファ11から読み出された2つのデータ D_1 、 D_2 は、波形窓掛け・加算部16、圧伸率制御部17及び出力バッファ18からなる圧伸処理制御手段に供給される。波形窓掛け・加算部16に供給されたデータ D_1 、 D_2 は、ここで所定の時間窓関数を乗算されて加算される。また、一方のデータ D_2 は、圧伸率制御部17にも供給されている。圧伸率制御部17は、制御部14から与えられる圧伸処理の対象長さ L の情報に基づいて原オーディオデータから波形を切り出す。圧伸処理の対象長さ L は、予め設定された圧伸率 R と、抽出された基本周期 L_p とに基づき制御部14で算出される。そして、波形窓掛け・加算部16で加算された波形と、圧伸率制御部17で切り出された原波形とが出力バッファ18において合成されて時間軸圧伸された出力信号 $y(t)$ が生成される。

【0014】次に、このように構成されたこの実施例の装置の動作を説明する。図3は、アタック検出部1におけるアタック検出処理の手順を示すフローチャートである。アタックの位置は、信号電力 Pow とその時間微分 S_{pw} より求めることができる。信号電力 Pow の計算は、図4に示すように、予め定めた信号電力計算時間 T_1 の信号について、予め定めた信号電力評価更新時間長 T_2 で順次更新しながら行う。本発明者の調査によると、アタック検出のための信号電力計算時間 T_1 としては、3ms程度、信号電力評価更新時間長 T_2 としては1ms程度が望ましい。従って、ここでは、 T_1 、 T_2 としてこの数値を使用する。

【0015】まず、入力信号を $x(t)$ とし、前のアタック位置を $PreAtk$ とする（S1）。3msecの入力信号 $x(t)$ から信号電力 Pow を下記数1により求める（S3）。

【0016】

【数1】 $Pow = \sqrt{\sum x(t)}$

【0017】求められた信号電力 Pow に対し、しきい値（この例では“1000”）による評価を行うが（S6）、“アタック”といっても立上りが急峻であるだけで、立ち下がりはかなりの持続時間を持つものも多い。従って、1つ前のフレームの信号電力 $PrePow$ との差分絶対値 Dpw を数2のように求め（S5）、

【0018】

【数2】 $Dpw = \text{abs}(PrePow - Pow)$

【0019】この差分 Dpw がしきい値を超える場合を検

出する(S7, S8)。このとき、信号中で平均電力AvePowの大きな部分と小さな部分で、そのしきい値を変更すると良い。何故なら、平均電力AvePowの大きな部分では、その中にアタックが存在した場合、差分Dpwの値が小さなものになってしまう。また、信号電力Powの小さな部分では、アタックの急激な立上りにより、差分Dpwの値は大きなものとなる。具体的には、電力の平方根、つまり元の信号の振幅スケールに対して差分の値を、例えば信号電力Powの大きな部分に対しては500(S7)、小さな部分に対しては1000を適用する(S8)。なお、平均信号電力AvePowの評価においても同じく1000を適用する(S6)。

【0020】このように計算された信号電力Powに対して、その時間微分Spwを数3のように求める(S4)。

【0021】

【数3】 $Spw = dPow/dt$

【0022】この際、本来のアタックよりも少し前の場所を検出するため、過去3つのフレームの信号電力を平均し、それを元に微分値を計算すると良い(実際には傾き計算)。この傾きがしきい値(例えば45度)以上の場合を検出する(S7, S8)。

【0023】以上の処理によりアタックの候補Atkが検出される。但し、本発明者の調査によると、アタックの間隔は殆どが30msec以上の間隔となっているため、アタックを検出した場合には、それが前回検出したアタックから30msec以上間隔を空けているかどうかを検出条件としている(S10, S11)。アタックが検出されなかったら、平均電力AvePow及び前回電力PrePowを更新して(S12)、以上の処理を繰り返す。アタックが300msecを超えても存在しない場合には、300msecを区切りとして時間軸圧伸する(S2, S13)。

【0024】いま、図5に示すように、8secと8.03secの位置でアタックが検出されたとし、伸長率が120%であるとする、アタック間の30msecの信号が36msecに伸長される。時間軸伸長後の出力信号y(t)の最初のアタック位置がそれまでの伸長処理により決定される位置、例えば9.6secであれば、次のアタック位置は、36msec後の9.636secとなる。

$$S = \frac{1}{L_p} \sum_{i=0}^{L_p-1} \{D(t_x) - D(t_x + L_p)\}^2$$

【0028】この場合、類似度Sが小さいほど類似性が高いことを示すことになる。勿論、これは一例であって、このような二乗誤差の他に、誤差の絶対値和や自己相関関数を用いることもできる。

【0029】ここで、例えば図9に示すように、アタック位置間の区間の前端部分(アタック位置)及び後端部分(次のアタック位置の直前の区間)の信号は、そのままとし、その中間部分の信号を時間軸圧伸処理する。時間軸圧伸処理は、その両端において、時間軸圧伸されな

【0025】次に、時間軸圧伸処理部2での時間軸圧伸処理について説明する。図6及び図7は、この時間軸圧伸手法を説明するための図であり、図6は圧縮処理、図7は伸長処理をそれぞれ示している。まず、同図(a)に示すように、原オーディオデータの時間軸方向の隣接波形区間の類似性判定処理を行って基本周期Lpを抽出する。具体的には、区間長の初期値を最小値Lminに設定して隣接する区間長Lminの波形の類似度を判定する。これを区間長が最大値Lmaxとなるまで繰り返し、最も類似していると判定された区間長を、同図(b)のように基本周期Lpと決定する。次に、決定された基本周期Lpの隣接する2つの波形に、同図(c)に示すような窓関数を掛けて、これらを同図(d), (e), (f)に示すように重ね合わせる。図6(f)のように、重ね合わせた波形を2つの基本周期の波形と置き換えれば時間軸圧縮となり、図7(f)のように、重ね合わせた波形を2つの基本周期の波形の間に挿入すれば時間軸伸長となる。

【0026】図8は、この時間軸圧伸処理の手順を示すフローチャートである。入力信号x(t)は、必要な量がまず遅延バッファ11に格納される(S21)。このバッファ11は、最低でも2×Lmaxサンプルの容量が必要である。次に、類似性判定のための区間長Lpの初期値として最小値Lminが与えられ、類似度Sとして最大値Smaxが与えられる(S22)。そして類似度Sが計算されると共に(S23)、区間長Lpを1つずつ増やして(S24)、Lpが最大値Lmaxに達するまで、類似度Sを計算して(S25, S23)、最も類似性の高かった区間長Lpを求める(S23)。ここで図6及び図7を参照して明らかなように、類似性判定は、現在点T0からT0+Lp-1間での区間の波形Wave Aと、T0+LpからT0+2Lpまでの区間の波形Wave Bとの類似度演算となる。いま、これら区間の対応する各時間軸方向の位置をtx, tx+Lpとすると、類似度Sは、下記数4のように二乗誤差によって求めることができる。

【0027】

【数4】

い信号と滑らかに結合されるように行う。最も目立つアタックの部分の波形は、そのまま維持されるので、本来の音に近い音が得られる。

【0030】ところで、このようにアタック位置を基本とする時間軸圧伸処理では、処理はアタック間で完結し、その前後の信号は一切用いないことが重要である。しかも、前述のように、時間軸圧伸処理された信号と時間軸圧伸処理されない信号とを滑らかに接続しなければならぬ。この場合、時間軸圧伸処理をポイント移動量

制御による重複加算法によって行くと、必ず処理しきれない部分が発生する。特に時間軸圧伸率が100%に近い部分では、この部分が非常に長くなる。

【0031】図10は、時間軸伸長時に処理しきれなかった部分をアタック位置間の後端部分からクロスフェードに必要な分のデータを取り出し、一部をクロスフェードして時間的な辻褄を合わせる処理を示している。図11は、時間軸伸長におけるクロスフェード時に、データが足りない場合に、一部のデータを繰り返して伸長を行っている様子を示している。これはポインタ移動量が大きすぎて全く処理できない場合等に有効である。

【0032】図12は、時間軸圧縮時の様子を示したもので、伸長時と同様に、処理しきれなかった部分をクロスフェードで時間軸圧縮している。圧縮時にはデータが不足することはあり得ないので、全てアタック位置間の後端部分から必要なデータを取り出し、クロスフェードすればよい。

【0033】以上の処理では、モノチャンネルのみの処理について説明したが、処理すべきリズム音源信号がステレオ信号である場合には、L、Rのそれぞれのチャンネルの信号に対して独立に処理を行うと、ステレオ再生した場合に、定位感が広がってしまうことがある。これは、上述した時間軸圧伸を行った際、左右チャンネル間でクロスフェードポイントがずれることにより、位相が変化してしまうため定位感が大きく損なわれるものと推測される。そこで、例えば図13に示すように、L+Rの信号に対してアタック検出及びポインタ評価処理(Lpの決定処理)を行うアタック検出部21及びポインタ処理部22を設けると共に、これらによって得られた左右チャンネル共通のアタック位置及び区間長Lpを使用して、各チャンネルを個別に処理する圧伸制御部23、24を設けるようにすれば、圧伸処理に伴う左右チャンネルの位相変化は最小限に抑えられ、本来の定位感を大きく損なうことが少なくなる。

【0034】

【発明の効果】以上述べたように、この発明によれば、リズム音源信号のアタック位置を検出し、検出されたアタック位置間で時間軸圧伸処理を施すようにしているの

で、二度打ちが発生することはなく、アタック位置の間隔も圧伸率に応じて圧縮又は伸長されるので、リズムの狂いが発生するようなこともない。

【図面の簡単な説明】

【図1】 この発明の一実施例に係るリズム音源信号の時間軸圧伸装置のブロック図である。

【図2】 同装置における時間軸圧伸処理部の構成を示すブロック図である。

【図3】 同装置におけるアタック検出部の処理を示すフローチャートである。

【図4】 同処理における信号電力計算時間と更新時間を説明するための図である。

【図5】 同装置による時間軸伸長処理前後の信号の様子を示す波形図である。

【図6】 同処理における時間軸圧縮処理を示す波形図である。

【図7】 同装置における時間軸伸長処理を示す波形図である。

【図8】 同装置における時間軸圧伸処理のフローチャートである。

【図9】 この発明における他の実施例に係る時間軸伸長処理前後の信号を示す波形図である。

【図10】 同処理におけるクロスフェード処理を説明するための図である。

【図11】 同処理におけるクロスフェード処理を説明するための図である。

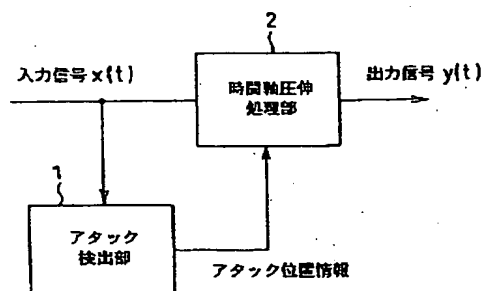
【図12】 同処理におけるクロスフェード処理を説明するための図である。

【図13】 この発明の更に他の実施例に係る時間軸圧伸装置のブロック図である。

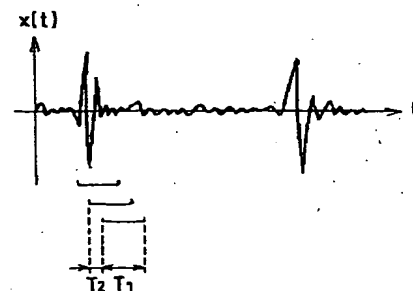
【符号の説明】

1、21…アタック検出部、2…時間軸圧伸処理部、11…遅延バッファ、12…隣接波形読出部、13…類似度演算部、14…制御部、15…波形読出制御部、16…波形窓掛け・加算部、17…圧伸率制御部、18…出力バッファ、22…ポインタ制御部、23、24…圧伸制御部。

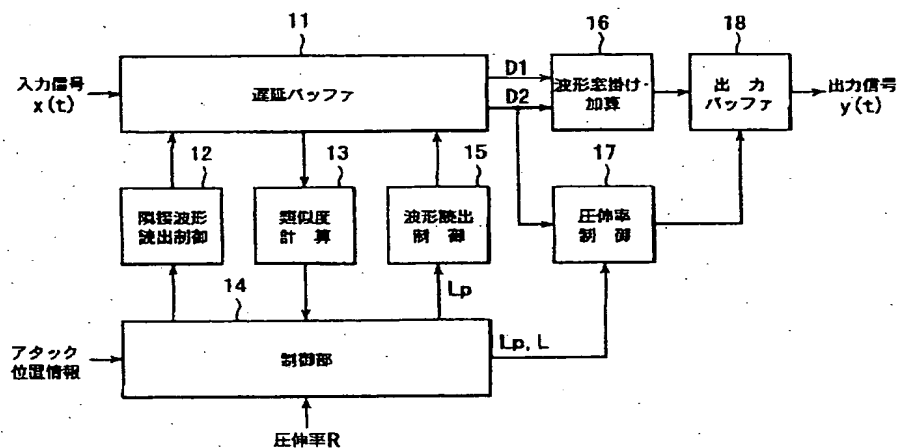
【図1】



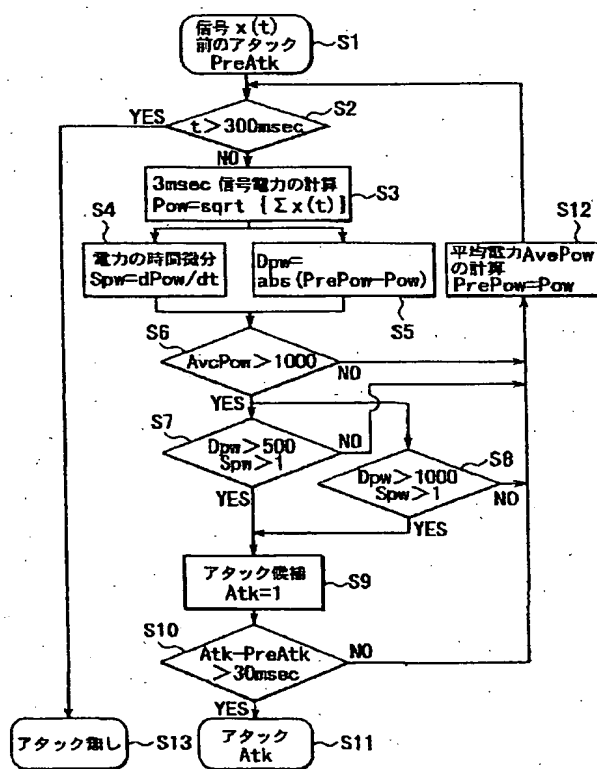
【図4】



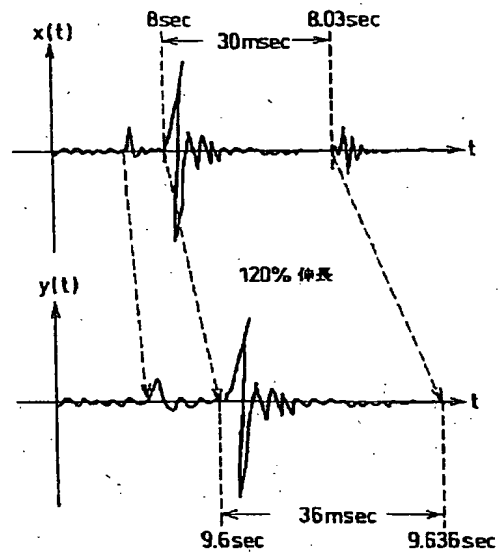
【図2】



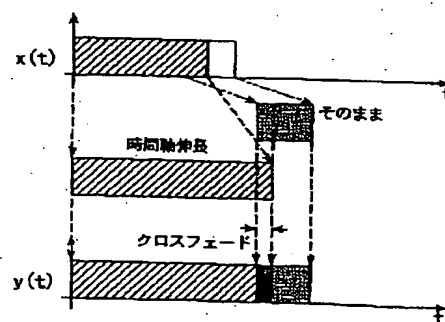
【図3】



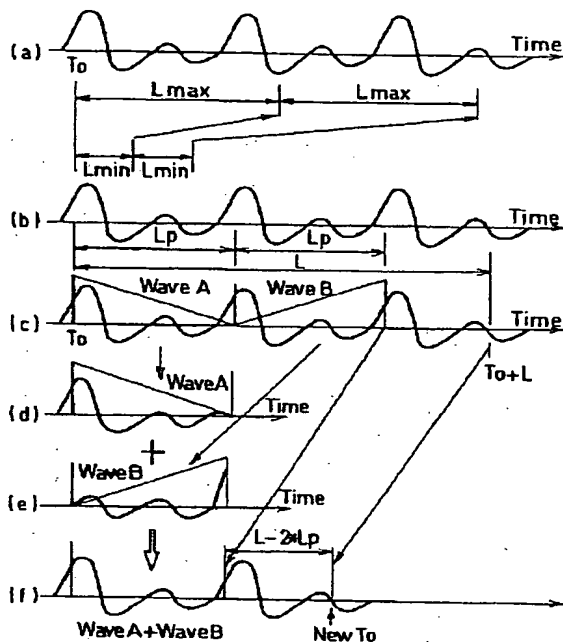
【図5】



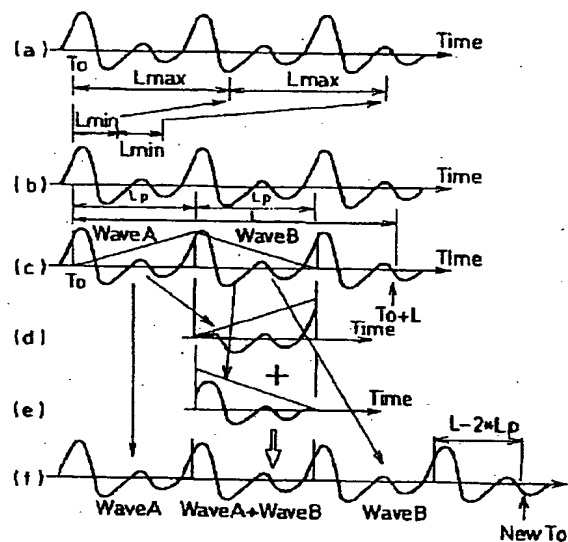
【図10】



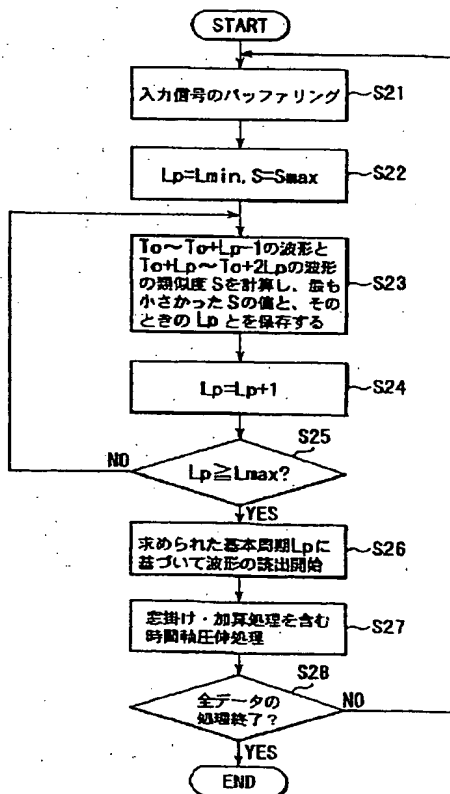
【図6】



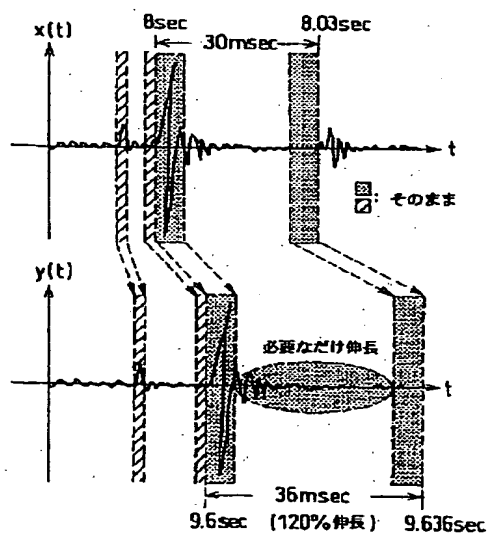
【図7】



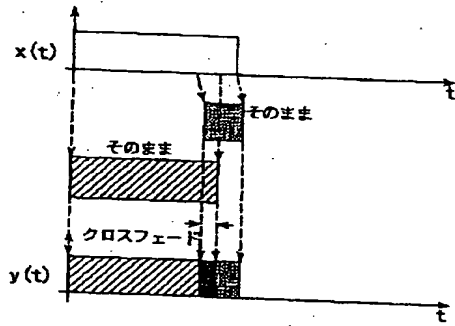
【図8】



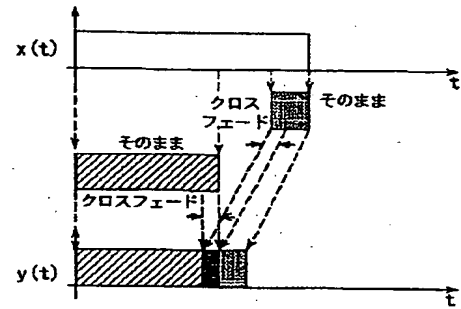
【図9】



【図11】



【図12】



【図13】

